

STUDENT ID: 1002088875

NAME: PRIYA S KULKARNI

Hierarchical Clustering for Seed Categorization Project:

1. Dataset: The dataset used in this study is publicly available on the UCI website (<http://archive.ics.uci.edu/ml/datasets/seeds>)
2. I have downloaded the dataset from the above link and divided the dataset into training and test dataset in 70:30 ratio respectively.
 - The training dataset is in the file seeddataset.csv.
 - The testing dataset is in the file seeddatasettest.csv.
3. The implementation of the project is in the file ML_Final_Project.ipynb .

Instruction to run the code: Google Colab was used as the execution environment. To ensure that the dataset is accessible to the code, it needs to be uploaded to Google Drive and its path should be specified in the code. Once the dataset is uploaded and the path is specified, the user can run all the cells in the code to see the output for each cell.

Steps to run the code:

- Open google colab.
- Upload the .csv data files respectively.
→The dataset files are in .csv files: seeddataset.csv and seeddatasettest.csv.
- Please place the path of the file in read_csv line respectively of the first cell in ML_Final_Project.ipynb.
→for example: `pd.read_csv('/content/drive/MyDrive/ML/seeddataset.csv',dtype='object')`
- Select "Runtime" on top menu and choose "Run all" to run all cells.
- The output is seen respectively.
- Comparison answers for required questions are written in the text cell.